Communication: Accurate description of interaction energies and three-body effects in weakly bound molecular complexes by PBE-QIDH models

A. J. Pérez-Jiménez, E. Brémond, C. Adamo, and J. C. Sancho-García

Citation: The Journal of Chemical Physics **149**, 041101 (2018); doi: 10.1063/1.5042153
View online: https://doi.org/10.1063/1.5042153
View Table of Contents: http://aip.scitation.org/toc/jcp/149/4
Published by the American Institute of Physics

Communication: Accurate description of interaction energies and three-body effects in weakly bound molecular complexes by PBE-QIDH models

A. J. Pérez-Jiménez,[1,a)] E. Brémond,[2] C. Adamo,[3,4] and J. C. Sancho-García[1,b)]
[1]*Department of Physical Chemistry, University of Alicante, E-03080 Alicante, Spain*
[2]*Université Paris Diderot, Sorbonne Paris Cité, ITODYS, UMR CNRS 7086, F-75013 Paris, France*
[3]*Chimie ParisTech, PSL Research University, CNRS, Institut de Recherche de Chimie Paris, F-75005 Paris, France*
[4]*Institut Universitaire de France, F-75005 Paris, France*

(Received 30 May 2018; accepted 12 July 2018; published online 25 July 2018)

We apply a recently developed parameter-free double-hybrid density functional belonging to the quadratic-integrand double-hybrid model to calculate association energies (ΔE) and three-body effects ($\Delta^3 E$) arising from intermolecular interactions in weakly bound supramolecular complexes (i.e., the dataset 3B-69). The model behaves very accurately for trimer association energies and is found to outperform widely used density functional approximations while approaching the accuracy of more costly *ab initio* methods for three-body effects. The results are further improved when we add some specific corrections for the remaining dispersion interactions, D3(BJ) or VV10 for two-body effects and Axilrod-Teller-Muto for three-body effects, leading to marginal deviations (less than 1 kcal/mol for ΔE and around 0.03–0.04 kcal/mol for $\Delta^3 E$) with respect to benchmark results. *Published by AIP Publishing.* https://doi.org/10.1063/1.5042153

The interaction energy of a nanoaggregate of N rigidly interacting molecules is calculated as

$$\Delta E = E(\ldots XYZ) - E(Z) - E(Y) - E(X) - \cdots, \quad (1)$$

where $E(M)$, $M = \ldots, X, Y, Z$, is the energy of individual molecules (i.e., monomers) and $E(\ldots XYZ)$ is the energy of the weakly bound nanoaggregate. Using the many-body decomposition approach, valid not only in terms of computational efficiency but also useful for providing physical insights, the energy can be decomposed as follows:

$$\Delta E = \sum_{X<Y} \Delta^2 E(XY) + \sum_{X<Y<Z} \Delta^3 E(XYZ) + \cdots \quad (2)$$

with Δ^n being the nth order term. The high-order terms, $n \geq 4$, are often neglected due to the fast convergence of the above expansion. The explicit form of the two-body $[\Delta^2 E(XY)]$ and three-body $[\Delta^3 E(XYZ)]$ terms is

$$\Delta^2 E(XY) = E(XY) - E(Y) - E(X), \quad (3)$$

$$\begin{aligned}\Delta^3 E(XYZ) &= E(XYZ) - \Delta^2 E(XY) - \Delta^2 E(XZ) - \Delta^2 E(YZ) \\ &\quad - E(Z) - E(Y) - E(X) \\ &= E(XYZ) - E(XY) - E(XZ) - E(YZ) \\ &\quad + E(Z) + E(Y) + E(X), \end{aligned} \quad (4)$$

after replacing in the latter expression the corresponding two-body specific terms. The energy of weakly bound trimers, $E(XYZ)$, and dimers, e.g., $E(XY)$, is needed, as well as those of monomers, e.g., $E(X)$, for calculating the three-body $\Delta^3 E(XYZ)$ term. Few datasets were developed to probe these properties, and, for that purpose, we choose the state-of-the-art 3B-69 dataset,[1] which is composed of non-covalently bound trimer geometries (×69) extracted from the reported crystalline structures of 23 compounds. The dataset comprises a variety of compounds (see the supplementary material for further details) and packing motifs, which in fact translates into intermolecular interactions of different nature (e.g., hydrogen bonds, polarization, and/or dispersion interactions) depending on each particular molecule and supramolecular configuration. The moderate size of the compounds allowed previous computations at the counterpoise-corrected coupled-cluster single double triple [CCSD(T)] complete basis set (CBS) level, and it becomes thus possible to benchmark any theoretical method and/or computational approach for three-body effects, whose examples are still scarce. Previously applied density functional approaches are indeed reported to perform poorly,[1,2] possibly due to the incomplete treatment of exchange and polarization effects even when the expressions were corrected for including dispersion energies approximately by resorting to interatomic pairwise potentials.

In this context, we will thus assess here the accuracy of the parameter-free Quadratic Integrand Double-Hybrid (QIDH) density functional,[3,4] PBE-QIDH, and its recently developed dispersion-corrected extensions,[5] PBE-QIDH-D3(BJ) and PBE-QIDH-VV10. This method linearly combines parameter-free exchange and correlation density functionals, $E_x[\rho]$ and $E_c[\rho]$, respectively, with a portion of EXact-like eXchange (EXX), $E_x^{\mathrm{EXX}}[\phi]$, and second-order Perturbation Theory (PT2), $E_c^{\mathrm{PT2}}[\phi, \phi']$, but with the

a)E-mail: aj.perez@ua.es
b)E-mail: jc.sancho@ua.es

corresponding weights ($\lambda_x = 3^{-1/3}$ and $\lambda_c = 1/3$) derived after imposing a set of first-principles constraints,

$$E_{xc}^{QIDH} = \lambda_x E_x^{\mathrm{EXX}}[\phi] + (1-\lambda_x)E_x[\rho] + \lambda_c E_c^{\mathrm{PT2}}[\phi,\phi'] + (1-\lambda_c)E_c[\rho]. \quad (5)$$

The application of this model to the computation of trimer interaction energies $\Delta E = E(XYZ) - E(Z) - E(Y) - E(X)$ of the 3B-69 dataset, with the very large aug-cc-pVQZ basis set to reduce basis set incompleteness errors, leads to a Mean Signed Error (MSE), Mean Unsigned Error (MUE), and Root Mean-Squared Error (RMSE) of 1.41, 1.59, and 1.89 kcal/mol with respect to CCSD(T)/CBS results, respectively (see Table S1 of the supplementary material for the specific values). These low errors agree with what one would expect from applications to other datasets composed only of dimer interaction energies,[6] usually in the margin of 1–2 kcal/mol for the best methods in use.[7] Only a residual number of trimers are predicted to be unbound (20c and 22b, $\Delta E > 0$), which are precisely those needing a larger contribution of dispersion corrections (*vide infra*).

We can now judge the effect of adding state-of-the-art corrections for dispersion effects to the latter PBE-QIDH form, i.e., through interatomic pairwise or non-local interactions, which are

$$E^{D3(BJ)} = -\sum_{n=6,8} s_n \sum_{B>A}^{atom\ pairs} \frac{C_n^{AB}}{R_{AB}^n + f_n(R_{AB}^0)}, \quad (6)$$

$$E^{VV10} = \int d\mathbf{r}\rho(\mathbf{r})\left[\frac{1}{2}\int d\mathbf{r}'\Phi(\mathbf{r},\mathbf{r}',b)\rho(\mathbf{r}') + \beta(b)\right], \quad (7)$$

providing the corresponding PBE-QIDH-D3(BJ) or PBE-QIDH-VV10 expressions.[5] Whereas the function in Eq. (6) depends[8] on the interatomic $R_{AB} = |R_A - R_B|$ distance, the nth-order interatomic dispersion coefficients C_n^{AB}, and the damping function $f_n = (a_1 R_{AB}^0 + a_2)^n$ with $R_{AB}^0 = \sqrt{\frac{C_8^{AB}}{C_6^{AB}}}$, Eq. (7) relies[9] on the kernel $\Phi(\mathbf{r},\mathbf{r}')$ coupling the electronic densities at two different spatial points, $\rho(\mathbf{r})$ and $\rho(\mathbf{r}')$. The latter expression was recently coupled successfully to other double-hybrid functionals.[10] The set of parameters entering into each model (s_6, s_8, a_1, and a_2 for $E^{D3(BJ)}$ and b for E^{VV10}) is also given for completeness in Table I. The application of these corrections to the calculation of trimer energies, ΔE, reduces significantly the MSE and MUE values to 0.05 and 0.64 kcal/mol at the PBE-QIDH-D3(BJ) level and to −0.30 and 0.61 kcal/mol at the PBE-QIDH-VV10 level, with the corresponding RMSE being now, respectively, 0.80 and 0.84 kcal/mol, and thus achieving in both cases the very stringent threshold MUE/RMSE below ±1 kcal/mol and showing concomitantly the high transferability of the parameters derived before[5] when applied out of the training (e.g., S130) set.[11]

Having obtained these encouraging results for trimer association energies, and before calculating three-body effects[12] (i.e., $\Delta^3 E$), we bracket the accuracy of the PBE-QIDH model with respect to both *ab initio* and density functional methods by introducing the three-body (and geometry-based) Axilrod-Teller-Muto (ATM) correction[13] given in the form

$$E^{ATM} = \sum_{C>B>A}^{atom\ triples} C_9^{ABC} \frac{(3\cos\theta_{AB}\cos\theta_{BC}\cos\theta_{AC}+1)}{(R_{AB}R_{BC}R_{AC})^3} \times f_n(\overline{R}_{ABC}), \quad (8)$$

depending on the $R_{AB} - R_{BC} - R_{AC}$ ($\theta_{AB} - \theta_{BC} - \theta_{AC}$) interatomic distances (angles), $\overline{R}_{ABC}$ is the geometric mean of the former, $C_9^{ABC} \approx -\sqrt{C_6^{AB}C_6^{BC}C_6^{AC}}$, and f_n is another damping function. This expression was previously applied to large supramolecular complexes,[14] where its influence is expected to contribute the most to association energies due to extended polarization effects, and is known to provide a very satisfactory agreement for bulk three-body molecular effects with respect to CCSD(T) results for systems of moderate size as those contained in the 3B-69 dataset.[15] We find that the previously calculated errors for ΔE are not significantly affected (MSE, MUE, and RMSE of 0.11, 0.70, and 0.86 kcal/mol at the PBE-QIDH-D3(BJ)+ATM level) after adding this correction. Figures S1 and S2 of the supplementary material show the impact of the whole dispersion energy for the ΔE values, for which the contribution to the total energy can be found separately, and how it systematically corrects all the values in the right direction.

We now apply the whole set of dispersion corrections to the calculation of three-body effects $\Delta^3 E$, which would also allow us to compare with previous estimates in the literature. Table II (see also Table S2 of the supplementary material for specific values) gathers the statistical error values (MSE, MUE, and RMSE) for different methods as a function of their formal scaling with the system size (N), for which we can

TABLE I. Summary of the parameters entering into the PBE-QIDH dispersion-corrected versions.

Method	s_6	a_1	s_8	a_2	b
PBE-QIDH-D3(BJ)	0.610	0.114	0.566	7.538	...
PBE-QIDH-VV10	...	...	...	...	14.2

TABLE II. Statistical errors (kcal/mol) of QIDH-based models, as compared with previous estimates, for the calculation of $\Delta^3 E$ three-body effects.

Method	MSE	MUE	RMSE	Scaling
PBE-D3(BJ)+ATM[a]	0.093	0.122	0.147	$O(N^3)$
BLYP-D3(BJ)+ATM[a]	−0.084	0.086	0.111	$O(N^3)$
PBE0-D3(BJ)+ATM[a]	0.043	0.068	0.080	$O(N^4)$
B3LYP-D3(BJ)+ATM[a]	−0.046	0.052	0.069	$O(N^4)$
BH-LYP-D3(BJ)+ATM[a]	−0.018	0.033	0.045	$O(N^4)$
PBE-QIDH	−0.022	0.035	0.047	$O(N^5)$
PBE-QIDH-D3(BJ)	−0.022	0.034	0.047	$O(N^5)$
PBE-QIDH-D3(BJ)+ATM	0.002	0.026	0.036	$O(N^5)$
PBE-QIDH-VV10	−0.032	0.041	0.054	$O(N^5)$
B2-PLYP-D3(BJ)+ATM[a]	−0.028	0.038	0.050	$O(N^5)$
MP2[a]	−0.039	0.045	0.059	$O(N^5)$
MP3[a]	0.022	0.026	0.035	$O(N^6)$
CCSD[a]	0.010	0.014	0.019	$O(N^6)$
SCS-CCSD[a]	−0.001	0.010	0.015	$O(N^6)$
MP2.5[a]	−0.009	0.014	0.019	$O(N^6)$

[a]Values taken from Ref. 1.

extract the following general conclusions: (i) going across the approximate hierarchy of density functionals (i.e., semilocal, hybrid, and double-hybrid) systematically reduces the errors, e.g., compare PBE, PBE0, and PBE-QIDH, or BLYP, B3LYP, and B2-PLYP results, as well as it does to include a larger value of λ_x in hybrid-only methods, e.g., compare B3LYP and BH-LYP values, which is related to the self-interaction error of commonly used functionals;[16] (ii) without including any correction for dispersion, the pristine PBE-QIDH model becomes the best density functional of all the set (PBE, PBE0, BLYP, B3LYP, BH-LYP, CAM-B3LYP, TPSS, M06-2X, M06-HF, B2-PLYP) tested up to now,[1] even when previous results from the literature explicitly incorporate two- and three-body dispersion corrections and/or some functionals (e.g., M06-2X) are obtained including weak interactions into their training set; and (iii) the role of the three-body ATM correction reduces even further the error obtained with the PBE-QIDH-D3(BJ) model, achieving an accuracy competitive with methods such as MP2.5 or (SCS-)CCSD which naturally incorporates three-body interactions arising from correlation effects. Note, however, that we neglect at this level other higher-order effects (i.e., fourth-order and beyond) and many-body non-additivity of dispersion energies.[17] Figure S3 of the supplementary material shows the impact of the ATM correction for the Δ^3E values, ranging in all cases between 95% and 105% and thus indicating its predominant role here. We also note the effectiveness of the MP2+ATM coupling for Δ^3E estimates,[2] which also retains the accuracy of MP2.5 or (SCS-)CCSD methods.

In summary, we show the high quality of the PBE-QIDH functional for describing interelectronic effects of all types arising from weakly bound interactions. The method performs well across the whole 3B-69 benchmark set, for low or highly polarizable systems, and prompts to reconsider previous findings about density functional approaches and their poor performance for three-body effects, where exchange and polarization contributions are of the same or higher importance than dispersion contributions.

See supplementary material for (i) the structures of all trimers belonging to the 3B-69 dataset; (ii) the computational details; (iii) all the ΔE individual values for the 3B-69 dataset at the PBE-QIDH, PBE-QIDH-D3(BJ), PBE-QIDH-VV10, and PBE-QIDH-D3(BJ)+ATM levels; (iv) the contribution of full and ATM dispersion energy to the final ΔE values at the PBE-QIDH-D3(BJ)+ATM level; (v) all the Δ^3E individual values for the 3B-69 dataset at the PBE-QIDH, PBE-QIDH-D3(BJ), PBE-QIDH-VV10, and PBE-QIDH-D3(BJ)+ATM levels; and (vi) the contribution of the ATM dispersion correction to the final Δ^3E values at the PBE-QIDH-D3(BJ)+ATM level.

We acknowledge the "Ministerio de Economía y Competitividad" of Spain and the "European Regional Development Fund" through the Project No. CTQ2014-55073-P.

[1] J. Rezac, Y. Huang, P. Hobza, and G. J. Beran, "Benchmark calculations of three-body intermolecular interactions and the performance of low-cost electronic structure methods," J. Chem. Theory Comput. **11**, 3065–3079 (2015).
[2] Y. Huang and G. J. Beran, "Reliable prediction of three-body intermolecular interactions using dispersion-corrected second-order Møller-Plesset perturbation theory," J. Chem. Phys. **143**, 044113 (2015).
[3] É. Brémond, J. C. Sancho-García, Á. J. Pérez-Jiménez, and C. Adamo, "Communication: Double-hybrid functionals from adiabatic-connection: The QIDH model," J. Chem. Phys. **141**, 031101 (2014).
[4] E. Bremond, I. Ciofini, J. C. Sancho-García, and C. Adamo, "Nonempirical double-hybrid functionals: An effective tool for chemists," Acc. Chem. Res. **49**, 1503–1513 (2016).
[5] J. C. Sancho-García, É. Brémond, M. Savarese, A. Pérez-Jiménez, and C. Adamo, "Partnering dispersion corrections with modern parameter-free double-hybrid density functionals," Phys. Chem. Chem. Phys. **19**, 13481–13487 (2017).
[6] P. Hobza, "Calculations on noncovalent interactions and databases of benchmark interaction energies," Acc. Chem. Res. **45**, 663–672 (2012).
[7] S. Grimme, A. Hansen, J. G. Brandenburg, and C. Bannwarth, "Dispersion-corrected mean-field electronic structure methods," Chem. Rev. **116**, 5105–5154 (2016).
[8] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate *ab initio* parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu," J. Chem. Phys. **132**, 154104 (2010).
[9] O. A. Vydrov and T. Van Voorhis, "Benchmark assessment of the accuracy of several van der Waals density functionals," J. Chem. Theory Comput. **8**, 1929–1934 (2012).
[10] J. Aragó, E. Ortí, and J. C. Sancho-García, "Nonlocal van der Waals approach merged with double-hybrid density functionals: Toward the accurate treatment of noncovalent interactions," J. Chem. Theory Comput. **9**, 3437–3443 (2013).
[11] L. Goerigk and S. Grimme, "Efficient and accurate double-hybrid-meta-GGA density functionals evaluation with the extended GMTKN30 database for general main group thermochemistry, kinetics, and noncovalent interactions," J. Chem. Theory Comput. **7**, 291–309 (2010).
[12] R. Bukowski and K. Szalewicz, "Complete *ab initio* three-body nonadditive potential in Monte Carlo simulations of vapor–liquid equilibria and pure phases of argon," J. Chem. Phys. **114**, 9518–9531 (2001).
[13] O. Anatole von Lilienfeld and A. Tkatchenko, "Two-and three-body interatomic dispersion energy contributions to binding in molecules and solids," J. Chem. Phys. **132**, 234109 (2010).
[14] J. Calbó, E. Ortí, J. C. Sancho-García, and J. Aragó, "Accurate treatment of large supramolecular complexes by double-hybrid density functionals coupled with nonlocal van der Waals corrections," J. Chem. Theory Comput. **11**, 932–939 (2015).
[15] M. R. Kennedy, A. R. McDonald, A. E. DePrince, M. S. Marshall, R. Podeszwa, and C. D. Sherrill, "Communication: Resolving the three-body contribution to the lattice energy of crystalline benzene: Benchmark results from coupled-cluster theory," J. Chem. Phys. **140**, 121104 (2014).
[16] G. J. Beran, "Modeling polymorphic molecular crystals with electronic structure theory," Chem. Rev. **116**, 5567–5613 (2016).
[17] A. M. Reilly and A. Tkatchenko, "van der Waals dispersion interactions in molecular materials: Beyond pairwise additivity," Chem. Sci. **6**, 3289–3301 (2015).